12. TECHNIQUES USED IN McDONNELL DOUGLAS SUBJECTIVE NOISE-RATING STUDIES

By Lawrence E. Langdon and Richard F. Gabriel

McDonnell Douglas Corporation

12

SUMMARY

Recordings of DC-8-55 flyovers with both treated and existing nacelle configurations are being used in a psychoacoustic test. The results will be used to provide an empirical evaluation of the subjective effects of the nacelle treatment and an evaluation of existing noise-rating scales in terms of their accuracy in predicting the results of this particular change in noise spectra.

Outdoor recordings were selected to represent all combinations of three nominal thrust levels (15 000, 10 000, and 5000 lb) and three heights (500, 1000, and 2500 ft) for both the treated and existing nacelle configurations. Indoor recordings of treated and existing nacelle configurations were selected for the three thrust conditions for one location. Pair-comparison tests in an anechoic chamber using 40 college-age subjects are currently in progress, all subjects having normal hearing. Pairs of noise test stimuli were constructed to provide four comparisons $\left(\pm 1\frac{1}{2} \text{ and } \pm 4 \text{ dB}\right)$ for each condition. Pairs were chosen to equate for order effects and existing or treated as standard. Pairs of noise stimuli are being presented to the subjects in a random order.

Results will be analyzed to provide a statistical statement concerning the accuracy of the prediction of various existing scales such as EPNL, PNL, dB(A), and others.

INTRODUCTION

The usefulness of a body of knowledge generally increases as a function of the degree that mathematical formulations may be validly applied. Most, if not all the time, this usefulness requires quantitative data in the form of scalar information. Measurement seems to come naturally to the physical sciences but it was only a relatively short time ago that what is now called an inch was the length of the first joint of a king's index finger. Naturally, this length varied from king to king; consequently, its value was limited. However, as soon as a standard, although arbitrary, unit was established, order could be and was established.

A scale by itself has little value but a valid scale increases the accuracy of prediction and allows one to quantify and generalize from one situation to another in a convenient

manner. A valid and reliable psychoacoustic scale of annoyance resulting from noise would be of significant value in aiding aircraft designers and operators to reduce airport noise. Such a scale is mandatory for the establishment of equitable noise levels.

During the past several decades, a number of noise-rating scales have been developed. (See table I.) Original work was primarily concerned with quantifying loudness. Later work attempted to account for other factors in annoyance, such as pure tones or duration of the sound. The introduction of jet aircraft and helicopters into general use has further complicated the situation since the resulting noises make it necessary to consider additional noise variables. Each major engine development suggests the need for a new scale because of significant changes in the spectrum.

The development of a valid psychoacoustic scale for an essentially aesthetic judgment is not easy. Among the factors which complicate this task are

(1) The many attributes of the stimulus

(2) The wide variance of the observer characteristics

(3) Consideration of environmental factors

(4) Difficulty in specifying a precise stimulus

(5) Difficulty of control, even if the stimulus is adequately specified

Table II lists some of the factors which may contribute to scaling error.

These factors do not negate the use of psychoacoustic scales, they simply suggest caution. The DC-8-55 nacelle treatment has resulted in an important change in engine noise spectra. The pure-tone content associated with compressors has been significantly reduced. It was therefore judged important to conduct a psychoacoustic study to determine which of the existing scales most accurately predicts the human response of rating the effect of the nacelle treatment. Furthermore, since house attenuation influences the high-frequency segment of the spectra, as does the nacelle treatment, it was considered advisable to evaluate the accuracy of predicting indoor as well as outdoor spectra.

METHOD

Forty college students with normal hearing (within ±5 dB of the American Medical Association standard) are being used as subjects. Approximately 20 of the subjects are female and the remainder are male.

Recordings of flyover noise were obtained during the DC-8-55 flyover test program described by Marsh (paper no. 5). Representative samples of these recordings were provided to Dr. K. Kryter of Stanford Research Institute who is performing the study. All combinations of the following outdoor conditions are being evaluated:

(1) Thrust: takeoff (15 000 lb), takeoff with cutback (10 000 lb), and landing (5000 lb)

(2) Height: 500 ft, 1000 ft, and 2500 ft

(3) Nacelle treatment: existing and treated

Recordings were obtained within a house located under the flight path for thrusts of 5000, 10 000, and 15 000 lb which corresponded to altitudes of 400, 2000, and 1500 ft, respectively. Both treated- and existing-nacelle recordings were obtained for these conditions.

These recordings were used to prepare test tapes which consisted of stimuli paired to provide a standard stimulus adjacent to a comparison stimulus for each condition noted above. This technique is commonly known as "pair comparisons," or the method of constant-stimulus differences. Treated- and existing-nacelle engine noises were used as both standard and comparison sounds. The standard sound was compared with each of four comparison sounds. The level of the comparison sounds was changed so the standard was bracketed by comparison stimuli at $\pm 1\frac{1}{2}$ and ± 4 dB. The duration of the sound was determined by the 15-dB-down points of the respective noises. Stimulus pairs were randomly ordered.

All subjects are rating all the sounds. Instructions are to rate the sounds in terms of annoyance "if heard 20 to 30 times a day in your home" and other similar words to provide a global mental set. Ten subjects at a time are rating noises in an anechoic chamber. Listening positions have been calibrated and a spectrum shaper has been used to insure an accurate representation of flyover spectra at the subjects' ears. The chamber is rectangular with two adjacent speakers on one wall to provide total coverage.

DATA ANALYSIS TECHNIQUES

The data were being collected as this paper was written. Thus, definitive results cannot be presented at this time. The methods, however, to be used in analyzing the data will be discussed.

This experiment is attempting to answer several questions. These questions may be expressed as follows:

(1) Which scale best predicts obtained values and how accurate is it?

(2) How accurate is EPNL, and is it significantly worse than the best scale?

(3) How much reduction in perceived noise actually was obtained as a result of the nacelle treatment under various operationally meaningful conditions?

(4) Are there differences between the accuracy of prediction for indoor data as opposed to that for outdoor data?

The error variance of each of the scales studied in predicting points of subjective equality will be calculated. Estimates of the accuracy of the experiment in predicting these variance values will then be determined. In addition, the means and standard deviations of the error scores will be calculated to provide insight into any existing scaling errors. It will then be possible to determine which scale proved most valid in predicting the test data and the accuracy to be expected from that scale. Graphs and charts will be provided to permit inspection of the data in order to determine any systematic or otherwise important sources of scaling error.

Thus, the most important immediate result will be a measure of the accuracy of EPNL data reported on the nacelle treatment in predicting annoyance ratings. Secondly, a judgment can be made about the value of presenting these results in terms of another scale. Finally, specific sources of error in prediction will be isolated.

DISCUSSION

It was noted in the Introduction that scales, though inaccurate, may be useful. It was also noted that the greater the accuracy, the greater their value. The methodology commonly used in annoyance scaling and represented in most aspects by this study provides a maximum amount of control of many potentially important, confounding elements. It is limited, however, by its artificiality and lack of consideration for many important variables existing in the actual situation. To be highly accurate, an annoyance scale must consider such factors as those outlined in table II. Emphasis to date has been devoted primarily to the stimulus characteristics.

It may be true that humans can generate a mental set which provides judgments which are almost perfectly correlated with those that would be obtained in the actual situation. Data exist which indicate that speech interference is most frequently cited as the most annoying characteristic of aircraft noise (ref. 1). Kryter (ref. 2) has conducted a study which suggests that annoyance ratings correlate well with speech interference. Research to date, however, is only suggestive.

CONCLUDING REMARKS

The ultimate application of scaling techniques is to produce a tool to be utilized for such purposes as land-use planning, aircraft design, and operational constraints. As such, it must be remembered that present scales are measures of relative annoyance value of noise as measured at the ear and do not contain any practical implementation

factors. Furthermore, there are a number of factors, such as those listed in table II, which may have important but as yet unmeasured effects upon the annoyance resulting from noise and which suggest the importance of continuing research. Nevertheless, the present scaling techniques can be considered state-of-the-art approximations of human response and can be utilized with reasonable confidence.

REFERENCES

1. Hazard, William R.: Community Reactions to Aircraft Noise – Public Reactions. Progress of NASA Research Relating to Noise Alleviation of Large Subsonic Jet Aircraft, NASA SP-189, 1968, pp. 661-672.

2. Kryter, K. D.; and Williams, Carl E.: Masking of Speech by Aircraft Noise. J. Acoust. Soc. Amer., vol. 39, no. 1, Jan. 1966, pp. 138-150.

TABLE I.- SOME EXISTING NOISE-RATING SCALES

Scale	Variables considered	Comment
dB(A)	Frequency and amplitude	Single-frequency weighting per scale for all amplitudes
dB(B)	Frequency and amplitude	
dB(C)	Frequency and amplitude	
dB(D)	Frequency and amplitude	
LL	Frequency and amplitude	Separate curves for several amplitudes
PNL	Frequency and amplitude	
PNLT	Frequency, amplitude, and tone	
EPNL	Frequency, amplitude, tone, and duration	

TABLE II.- POTENTIAL FACTORS INVOLVED IN ANNOYANCE RATINGS

Stimulus characteristics	Amplitude Frequency (Hz) Duration Frequency of occurrence Change over time
Environmental characteristics	Ambient noise General nature of neighborhood Characteristics of the structure (attenuation) Weather conditions Others
Observer characteristics	Socioeconomic class Adaptation level Personal interests Tasks content Physiological characteristics
Other	Community attitudes Municipal legislation